Statistical Bias of Conic Fitting and Renormalization

Kenichi Kanatani

***Abstract*— Introducing a statistical model of noise in terms of the covariance matrix of the N-vector, we point out that the least-squares conic fitting is statistically biased. We present a new fitting scheme called renormalization for computing an unbiased estimate by automatically adjusting to noise. Relationships to existing methods are discussed, and our method is tested using real and synthetic data.**

***Index Terms*—Conic, ellipse, curve fitting, error analysis, renormalization.**

I. Introduction

If a robot is to operate in an industrial environment (say, in a nuclear power station), it must recognize gauges, meters, dials, and handles, most of which are circular; circles are perspectively projected into ellipses. Hence ellipses, or *conics* in general, are widely recognized in the study of computer vision as one of the most fundamental features. Detected conics provide more than clues to object recognition; if conics in an image are known to be perspective projections of conics in the scene of known shapes, their 3-D geometry can be computed analytically [7], [12], [17], [19], [20].

In order to do such analysis, image conics must be mathematically represented by curve fitting, and many studies of conic fitting have been done in the past [1]–[6], [8], [13]–[16], [18], [21]. The goal of such a study has customarily been thought of as finding a conic that passes near *observed data points* as closely as possible. We will point out in this correspondence that *this is not so*; the goal is to find a conic that passes through the (unknown) *true data points* from which observed data points are deviated. These two goals are identical for line fitting but not so for conics, due to the nonlinearity of the conic equation. Essentially, conic fitting is a *statistical inference* of inferring the locations of the true data points.

In this correspondence, we introduce a statistical model of noise and point out that the least-squares fit is *statistically biased* even if the weights are optimally chosen. We present a new fitting scheme called *renormalization* for computing an *unbiased* estimate by automatically adjusting to noise; we need not know noise characteristics. We will also discuss relationships to existing methods and test our method using real and synthetic data [15].

II. Least-Squares Conic Fitting

A *conic* is a quadratic curve on an xy plane expressed in the form

$$Ax^2 + 2Bxy + Cy^2 + 2(Dx + Ey) + F = 0. \tag{1}$$

Define a Z-axis perpendicular to the image passing through the image coordinate origin o. Let O be the point on the Z-axis at distance f from the image origin, and define X- and Y-axes at O in such a way that they are, respectively, parallel to the x- and y-axes (see Fig. 1). We call the origin O the *viewpoint* and the constant f the *focal length* on the analogy of the ideal camera model, though this model is hypothetical and adopted for the convenience of analysis.

Manuscript received October 21, 1991; revised December 7, 1992. Recommended for acceptance by Associate Editor R. Bolle.

The author is with the Department of Computer Science, Gunma University, Gunma 376, Japan.

IEEE Log Number 9212236.

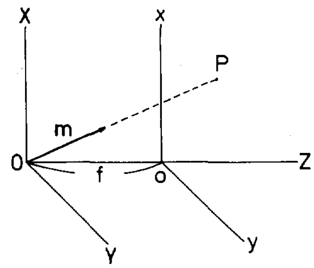


Fig. 1 The N-vector representing a point in the image.

With this setup, a point (x, y) in the image is uniquely represented by the unit vector m, starting from the viewpoint O and pointing toward it. It is easy to see that

$$m = N\left[\begin{pmatrix} x \\ y \\ f \end{pmatrix}\right], \tag{2}$$

where $N[\cdot]$ denotes normalization into a unit vector. Let us call the vector m the *N-vector* of point (x, y) [9]. Representing points by their N-vectors is equivalent to using (normalized) *homogeneous coordinates.* If we define matrix

$$Q = \begin{pmatrix} A & B & D/f \\ B & C & E/f \\ D/f & E/f & F/f^2 \end{pmatrix}, \tag{3}$$

(1) is written in terms of the N-vector (2) in the form

$$(m, Q, m) = 0, \tag{4}$$

where $(\cdot, \cdot)$ denotes the inner product of vectors. We call the conic represented by matrix Q simply "conic Q." Since any multiple of Q by a constant defines the same conic, we can adopt the normalization

$$\|Q\| = 1, \tag{5}$$

where $\|\cdot\|$ denotes the Euclidean matrix norm: $\|Q\|^2 = \sum_{i,j=1}^{3} Q_{ij}{}^2$.

Let $\{P_\alpha\}$, $\alpha = 1, \cdots, N$, be given data points, to which a conic is to be fitted. Let m_α be their N-vectors. Consider the least-squares optimization in the form

$$J(Q) = \sum_{\alpha=1}^{N} W_\alpha (m_\alpha, Qm_\alpha)^2 \to \min, \tag{6}$$

where W_α is the weight of the αth datum. Define the *moment tensor* $\mathcal{M} = (M_{ijkl})$ by

$$M_{ijkl} = \sum_{\alpha=1}^{N} W_\alpha m_{\alpha(i)} m_{\alpha(j)} m_{\alpha(k)} m_{\alpha(l)}, \tag{7}$$

where $m_{\alpha(i)}$ is the ith component of m_α. Tensor $\mathcal{M}$ determines a linear mapping from a matrix to a matrix: $\mathcal{M}Q$ is the matrix whose (ij) element is $\sum_{k,l=1}^{3} M_{ijkl}Q_{kl}$. If the matrix inner product is defined by $(A, B) = \sum_{i,j=1}^{3} A_{ij}B_{ij}$, the minimization (6) with the normalization (5) reads

$$J(Q) = (Q, \mathcal{M}Q) \to \min, \qquad \|Q\| = 1. \tag{8}$$

Since Q is a symmetric matrix, it can be identified with a six-dimensional vector

$$(Q_{11}, \sqrt{2}Q_{12}, Q_{22}, \sqrt{2}Q_{13}, \sqrt{2}Q_{23}, Q_{33})^\top. \tag{9}$$

Then, $\|Q\|$ can be viewed as either the matrix norm of matrix Q or the vector norm of vector Q. If tensor $\mathcal{M}$ is identified with the six-dimensional matrix at the bottom of this page, $(Q, \mathcal{M}Q)$ can be viewed as either the matrix inner product between matrixes Q and $\mathcal{M}Q$ or the vector inner product between vector Q and vector $\mathcal{M}Q$, the latter being regarded as vector Q multiplied by matrix $\mathcal{M}$.

If Q and $\mathcal{M}$ are identified with a vector and a matrix, respectively, the minimum of (8) is attained by the unit eigenvector Q of matrix $\mathcal{M}$ for the smallest eigenvalue. Alternatively, we can say that the minimum is attained by the unit "eigenmatrix" Q of tensor $\mathcal{M}$ for the smallest eigenvalue, where we mean by a unit "eigenmatrix" of tensor $\mathcal{M}$ for eigenvalue λ a matrix Q of norm 1 ($\|Q\| = 1$) such that $\mathcal{M}Q = \lambda Q$.

III. Statistical Model of Noise and Optimal Weights

Let $\bar{m}$ be the N-vector of a point in the image when there is no noise. In the presence of noise, a perturbed N-vector $m = \bar{m} + \Delta m$ is observed, where we regard inaccuracy of N-vectors computed from image data as caused by noise. We treat the noise Δm as a random variable. Consider the *covariance matrix*

$$V[m] = E[\Delta m \Delta m^{\top}], \tag{11}$$

where $\top$ denotes transpose and $E[\cdot]$ denotes the expectation over the statistical ensemble. We adopt the following model of noise: "Noise occurs at each data point P_α independently in the image and is equally likely in all orientations with the same root mean square ϵ." We call ϵ (measured in pixels) the *image accuracy*. If the size of the image is small compared with the focal length f, which can be chosen arbitrarily since the perspective camera model is hypothetical, the covariance matrix $V[m_\alpha]$ of the N-vector m_α has the form

$$V[m_\alpha] = \frac{\tilde{\epsilon}^2}{2}(I - \bar{m}_\alpha \bar{m}_\alpha^{\top}), \tag{12}$$

where I is the unit matrix and $\tilde{\epsilon} = \sqrt{E[\|\Delta m_\alpha\|^2]} = \epsilon / f$ [11].

It can be shown [10] that the most reasonable choice of the weights W_α of the least-squares estimation (2) in the sense of maximum likelihood estimation is

$$W_\alpha = \frac{\text{const.}}{\|Q\bar{m}_\alpha\|^2}. \tag{13}$$

Multiplication by a constant does not affect the solution, so we adopt the scaling $\sum_{\alpha=1}^{N} W_\alpha = 1$. Let us call the weights thus defined the *optimal weights*.

The optimal weights (13) convert the "algebraic distance" $|(m_\alpha, Qm_\alpha)|$ of the αth point to the conic Q into the "statistical distance" in the noise space, which can be identified with the "geometric distance" in the image to a first approximation according to our statistical model. The use of the weights of (13) corresponds to the methods discussed by Sampson [21], Bolle and Vemuri [2], and Safaee-Rad *et al.* [18]. The method proposed by Bookstein [3] corresponds to the use of uniform weights. He argues that the normalization $\|Q\| = 1$ is inappropriate because invariance under image translation is not assured. However, this argument is valid only when the algebraic distance is used; if the optimal weights are used, the choice of normalization becomes irrelevant.

IV. Statistical Bias and Unbiased Estimation

Let us denote the exact conic by $\bar{Q}$ to distinguish it from variable Q. The exact conic $\bar{Q}$ minimizes the unperturbed function $\bar{J}(Q) = (Q, \bar{\mathcal{M}}Q)$, where $\bar{\mathcal{M}}$ is the moment tensor for the exact data $\{\bar{m}_\alpha\}$, but the actual fitting minimizes the perturbed function $J(Q) = (Q, \mathcal{M}Q)$, and $\bar{Q}$ does not necessarily minimize $J(Q)$.

For the perturbed moment tensor $\mathcal{M} = \bar{\mathcal{M}} + \Delta\mathcal{M}$, we obtain the following proposition [10], where δ_{ij} is the Kronecker delta and $\tilde{\mu}^4 = E[\|\Delta m_\alpha\|^4]$:

Proposition 1:

$$\begin{aligned} E[\Delta M_{ijkl}] = \sum_{\alpha=1}^{N} W_\alpha \Big(& -3\Big(\tilde{\epsilon}^2 - \frac{\tilde{\mu}^4}{8}\Big)\bar{m}_{\alpha(i)}\bar{m}_{\alpha(j)}\bar{m}_{\alpha(k)}\bar{m}_{\alpha(l)} \\ & + \frac{1}{2}(\tilde{\epsilon}^2 - \frac{\tilde{\mu}^4}{4})(\delta_{ij}\bar{m}_{\alpha(k)}\bar{m}_{\alpha(l)} \\ & + \delta_{ik}\bar{m}_{\alpha(j)}\bar{m}_{\alpha(l)} + \delta_{il}\bar{m}_{\alpha(k)}\bar{m}_{\alpha(j)} \\ & + \delta_{jk}\bar{m}_{\alpha(i)}\bar{m}_{\alpha(l)} + \delta_{jl}\bar{m}_{\alpha(i)}\bar{m}_{\alpha(k)} \\ & + \delta_{kl}\bar{m}_{\alpha(i)}\bar{m}_{\alpha(j)}) \\ & + \frac{\tilde{\mu}^4}{8}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\Big). \end{aligned} \tag{14}$$

Suppose $J(Q)$ is minimized by $Q = \bar{Q} + \Delta Q$. Since Q is the eigenmatrix of $\mathcal{M} = \bar{\mathcal{M}} + \Delta\mathcal{M}$ for the smallest eigenvalue, the *perturbation theorem* tells us that the error ΔQ is linear in $\Delta\mathcal{M}$. This means that $E[\Delta Q]$ is linear in $E[\Delta\mathcal{M}]$. Hence, Q is *statistically biased*. According to the perturbation theorem, the statistical bias can be expressed explicitly. Let $\{Q^{(\kappa)}\}$, $\kappa = 0, \cdots, 5$, be the orthonormal system of eigenmatrices of $\bar{\mathcal{M}}$ for eigenvalues λ_κ ($\lambda_0 = 0$, $\lambda_\kappa > 0$, $\kappa = 1, \cdots, 5$). Let $\bar{M} = \sum_{\alpha=1}^{N} W_\alpha \bar{m}_\alpha \bar{m}_\alpha^{\top}$. Then, we obtain the following theorem [10]:

Proposition 2: To a first approximation, the optimally fitted conic Q is statistically biased by

$$E[\Delta Q] = -\sum_{\kappa=1}^{5}\Big(\frac{(4\tilde{\epsilon}^2 - \tilde{\mu}^4)\lambda_\kappa + \tilde{\mu}^4}{8\lambda_\kappa}U^{(\kappa)}Q + \frac{4\tilde{\epsilon}^2 - \tilde{\mu}^4}{2\lambda_\kappa}(U^{(\kappa)}\bar{M}Q)\Big)U^{(\kappa)}. \tag{15}$$

We now consider how to remove the above statistical bias. It appears that all we need to do is subtract from the computed conic the above bias. However, the bias theoretically derived above is expressed in terms of the "exact" N-vectors $\{\bar{m}_\alpha\}$ and the "exact" conic $\bar{Q}$; replacing them by perturbed m_α and Q would introduce another source of bias. It can be shown that the bias is removed by the following procedure [10].

Proposition 3: The unit eigenmatrix $\hat{Q}$ for the smallest eigenvalue of tensor $\hat{\mathcal{M}} = (\hat{M}_{ijkl})$ defined by

$$\hat{M}_{ijkl} = \sum_{\alpha=1}^{N} W_\alpha \Big(\Big(1 - \frac{\tilde{\epsilon}^2}{2}\Big)m_{\alpha(i)}m_{\alpha(j)}m_{\alpha(k)}m_{\alpha(l)}$$

$$\begin{pmatrix} M_{1111} & \sqrt{2}M_{1112} & M_{1122} & \sqrt{2}M_{1113} & M_{1123} & M_{1133} \\ \sqrt{2}M_{1211} & 2M_{1212} & \sqrt{2}M_{1222} & 2M_{1213} & 2M_{1223} & \sqrt{2}M_{1233} \\ M_{2211} & \sqrt{2}M_{2212} & M_{2222} & \sqrt{2}M_{2213} & \sqrt{2}M_{2223} & M_{2233} \\ \sqrt{2}M_{1311} & 2M_{1312} & \sqrt{2}M_{1322} & 2M_{1313} & 2M_{1323} & \sqrt{2}M_{1333} \\ \sqrt{2}M_{2311} & 2M_{2312} & \sqrt{2}M_{2322} & 2M_{2313} & 2M_{2323} & \sqrt{2}M_{2333} \\ M_{3311} & \sqrt{2}M_{3312} & M_{3322} & \sqrt{2}M_{3313} & \sqrt{2}M_{3323} & M_{3333} \end{pmatrix} \tag{10}$$

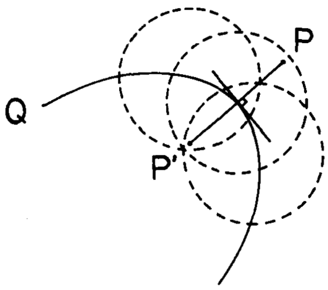


Fig. 2. Point P' is more likely to be observed than point P.

$$-\frac{1}{2}(\tilde{\epsilon}^2-\frac{\tilde{\mu}^4}{4})(\delta_{ij}m_{\alpha(k)}m_{\alpha(l)}+\delta_{ik}m_{\alpha(j)}m_{\alpha(l)}$$
$$+\delta_{il}m_{\alpha(k)}m_{\alpha(j)}+\delta_{jk}m_{\alpha(i)}m_{\alpha(l)}+\delta_{jl}m_{\alpha(i)}m_{\alpha(k)}$$
$$+\delta_{kl}m_{\alpha(i)}m_{\alpha(j)})+\frac{1}{2}\left(\tilde{\epsilon}^4-\frac{(\tilde{\epsilon}^2+2)\tilde{\mu}^4}{8}\right)$$
$$\cdot(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})) \quad (16)$$

is an unbiased estimate of Q to a first approximation.

What we have shown above is intuitively easy to understand in terms of the original equations (1) and (3). Matrix Q can be regarded as a vector consisting of A, $\sqrt{2}B$, C, $\sqrt{2}D/f$, $\sqrt{2}E/f$, and F/f^2. If $\{(x_\alpha, y_\alpha)\}$ are the data points, the quadratic form $(Q, \mathcal{M}Q)$ can be identified with

$$\sum_{\alpha=1}^{N} W_\alpha(Ax_\alpha{}^2+2Bx_\alpha y_\alpha+Cy_\alpha{}^2+2(Dx_\alpha+Ey_\alpha)+F)^2, \quad (17)$$

which is a fourth-degree polynomial in x_α and y_α. Consider the term $x_\alpha{}^2$, for example. If x_α is a random variable with mean $\bar{x}_\alpha$ and variance σ^2, we have $E[x_\alpha{}^2]=\bar{x}_\alpha{}^2+\sigma^2$. What we have pointed out is equivalent to asserting that the term $x_\alpha{}^2$, etc., in (17) should be replaced by $x_\alpha{}^2-\sigma^2$, etc., so that their expectations coincide with the exact $\bar{x}_\alpha{}^2$, etc.

This is a consequence of the fact that the curve is nonlinear. Many authors considered the (normal) distance to the curve to be a good measure of fit [2], [13], [18], [21], but this "geometric distance" does not completely agree with the "statistical distance." Consider Fig. 2, for example. Points P and P' are at the same distance from the curve. However, the fit of P is much worse than that of P'. The reason is as follows. Since the curve extends away from point P in both directions, it is unlikely that a point that was originally somewhere on the curve is displaced into the position of P by noise. On the other hand, it is much more likely that a point somewhere on the curve is displaced into the position of P', since the curve surrounds it. Thus, the geometric distance is not a good measure; we must consider the overall statistical effect of noise very carefully.

V. Renormalization

In order to compute an unbiased estimate using Proposition 3, we must know noise characteristics—the second and fourth moments $\tilde{\epsilon}^2$ and $\tilde{\mu}^4$, in particular. However, noise characteristics are different from image to image; they are difficult to predict *a priori* in real environments. In the following, we present an iteration scheme that automatically adjusts to noise; *we need not know noise characteristics.* We call this scheme *renormalization.*

Renormalization$(\{m_\alpha\})$:

1) Let $c=0$ and $Q=-I/\sqrt{3}$.
2) Compute

$$W_\alpha=\frac{1/\|Qm_\alpha\|^2}{\sum_{\beta=1}^{N}1/\|Qm_\beta\|^2}. \quad (18)$$

3) Compute matrix $M=(M_{ij})$ and tensor $\hat{\mathcal{M}}=(\hat{M}_{ijkl})$ by

$$M_{ij}=\sum_{\alpha=1}^{N}W_\alpha m_{\alpha(i)}m_{\alpha(j)}, \quad (19)$$

$$\hat{M}_{ijkl}=\sum_{\alpha=1}^{N}W_\alpha\big(m_{\alpha(i)}m_{\alpha(j)}m_{\alpha(k)}m_{\alpha(l)}$$
$$-\frac{c}{2}(\delta_{ij}m_{\alpha(k)}m_{\alpha(l)}+\delta_{ik}m_{\alpha(j)}m_{\alpha(l)}$$
$$+\delta_{il}m_{\alpha(k)}m_{\alpha(j)}+\delta_{jk}m_{\alpha(i)}m_{\alpha(l)}$$
$$+\delta_{jl}m_{\alpha(i)}m_{\alpha(k)}+\delta_{kl}m_{\alpha(i)}m_{\alpha(j)})\big). \quad (20)$$

Let Q be the unit eigenmatrix of $\hat{\mathcal{M}}$ for the smallest eigenvalue, and let λ_m be the smallest eigenvalue. Update c by

$$c\leftarrow c+\frac{\lambda_m}{Q(MQ)+(MQ^2)}. \quad (21)$$

Return Q if the update has converged; else go back to Step 2.

Example: Fig. 3(a) is an edge image obtained from a 300 × 200-pixel real image (f is set to 1000 pixels). Fig. 3(b) shows conic fitting to five ellipses by least squares with optimal weights. Fig. 3(c) is the result obtained by applying renormalization. We see that renormalization produces somewhat better fits, but the differences are small. In order to magnify the differences, the original image is cut in half (see Fig. 4(a)). Fig. 4(b) shows the fits obtained by least squares with optimal weights; Fig. 4(b) shows the fits obtained by renormalization. The fits are less accurate than those in Fig. 4(a). This is inevitable: All conic fitting schemes are bound to fail when obtained edge segments cover only a small portion of the entire conic. Yet, as we see, renormalization produces better results than the optimal least-squares scheme.

Fig. 5(a) shows 19 points on the upper half of $x^2+4y^2=1$. Each point is independently displaced by random noise obeying a two-dimensional Gaussian distribution with variance σ in x and y orientations, and we set $f=10$. The fourth-order moment $\tilde{\mu}^4$ can be expressed in terms of σ [10]. Fig. 5(b) shows the theoretical expectation of the conic predicted by Proposition 2 for $\sigma=0.01, 0.02, 0.03$. The exact conic is indicated by a broken line. We can see that conics fitted to such points are predicted to be flattened with larger eccentricities.

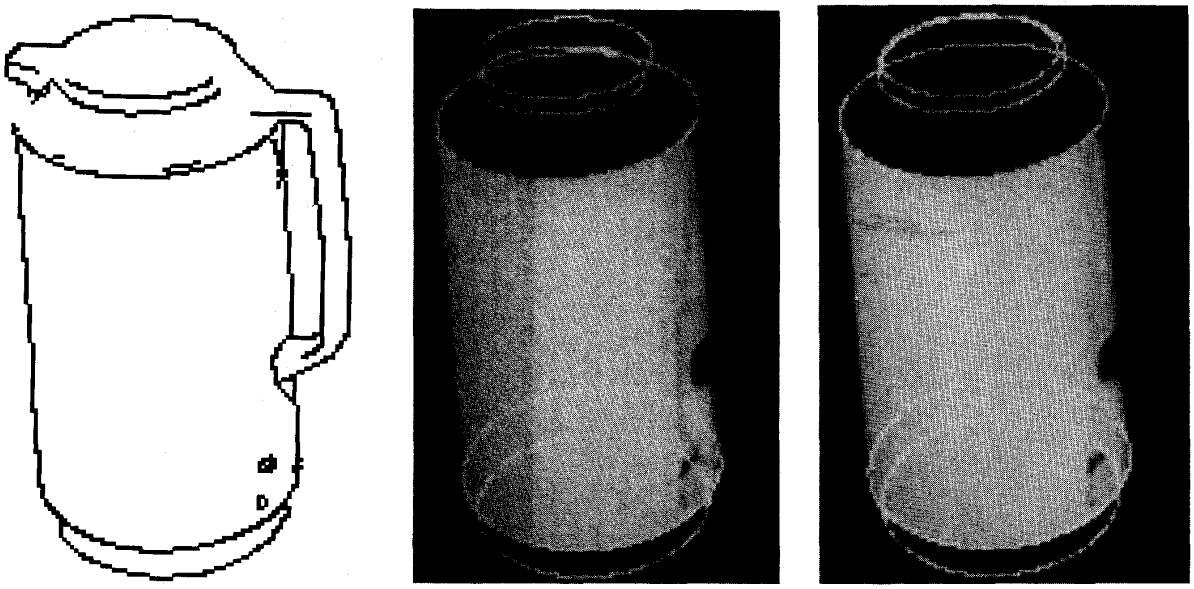
Fig. 6(a) shows 10 randomly chosen samples for $\sigma=0.02$ without renormalization. We can clearly observe the statistical bias predicted by Proposition 2. Fig. 6(b) shows 100 samples plotted with respect to the area S and the eccentricity e (the broken lines indicate the exact values $S=1.571$ and $e=0.866$). The sample average is $(\bar{S},\bar{e})=(1.367, 0.894)$. Fig. 7 shows the corresponding result with renormalization. It is clearly seen that the statistical bias has been removed. The sample average is $(\bar{S},\bar{e})=(1.608, 0.860)$.

In all the above examples, renormalization converges after three or four iterations.

VI. Convergence of Renormalization

The renormalization procedure described in the preceding section was obtained by the following reasoning. Ignoring the fourth-order quantities $\tilde{\epsilon}^4$ and $\tilde{\mu}^4$, and putting $c=\tilde{\epsilon}^2$, we can express the moment tensor $\hat{\mathcal{M}}=(\hat{M}_{ijkl})$ in the form

$$\hat{M}_{ijkl}=(1-\frac{c}{2})\sum_{\alpha=1}^{N}W_\alpha\big(m_{\alpha(i)}m_{\alpha(j)}m_{\alpha(k)}m_{\alpha(l)}$$
$$-\frac{c/2}{1-c/2}(\delta_{ij}m_{\alpha(k)}m_{\alpha(l)}+\delta_{ik}m_{\alpha(j)}m_{\alpha(l)}$$
$$+\delta_{il}m_{\alpha(k)}m_{\alpha(j)}+\delta_{jk}m_{\alpha(i)}m_{\alpha(l)}$$
$$+\delta_{jl}m_{\alpha(i)}m_{\alpha(k)}+\delta_{kl}m_{\alpha(i)}m_{\alpha(j)})\big). \quad (22)$$



(a) (b) (c)

Fig. 3. (a) An edge image. (b) Least-squares fitting. (c) Renormalization.

Ideally, the constant c should be determined so that $E[\hat{\mathcal{M}}] = \bar{\mathcal{M}}$, but this is impossible unless statistical noise characteristics are known. On the other hand, if $E[\hat{\mathcal{M}}] = E[\bar{\mathcal{M}}]$, we can prove the following [10]:

$$\begin{aligned} E[(\bar{Q}, \hat{\mathcal{M}}\bar{Q})] &= (\bar{Q}, E[\hat{\mathcal{M}}]\bar{Q}) \\ &= (1 - \frac{\tilde{\epsilon}^2}{2})(1 - 3(\tilde{\epsilon}^2 - \frac{\tilde{\mu}^4}{8}))(\bar{Q}, \bar{\mathcal{M}}\bar{Q}) \\ &= 0. \end{aligned} \quad (23)$$

Note that from (6) $J = (Q, \mathcal{M}Q)$ takes its absolute minimum 0 for the exact solution $\bar{Q}$ in the absence of noise. Hence, it is reasonable to choose c so that $(Q, \hat{\mathcal{M}}Q) = 0$ in each iteration step. Since the constant multiplier $1 - \hat{c}/2$ does not affect the eigenmatrix of $\hat{\mathcal{M}}$, it can be dropped from the above definition. Renaming $(c/2)/(1-c/2)$ as $c/2$, we obtain from (22)

$$(Q, \hat{\mathcal{M}}Q) = (Q, \mathcal{M}Q) - c(Qr(MQ) + 2(MQ^2)) \quad (24)$$

where M is the moment matrix of $\{m_\alpha\}$ defined by (5). Hence, if $(Q, \hat{\mathcal{M}}Q) \neq 0$ for the current estimates c and Q, then

$$(Q, \hat{\mathcal{M}}Q) - c'(Q(MQ) + 2(MQ^2)) = 0 \quad (25)$$

for

$$c' = \frac{(Q, \hat{\mathcal{M}}Q)}{Q(MQ) + 2(MQ^2)}. \quad (26)$$

Note that $(Q, \hat{\mathcal{M}}Q)$ is the smallest eigenvalue of $\hat{\mathcal{M}}$. From this observation, we obtain the procedure for renormalization given in the preceding section.

The procedure for renormalization can be expressed in an abstract form as follows. What we want to compute is the unit eigenvector $\bar{u}_m$ of a positive semidefinite matrix $\bar{A}$ for eigenvalue 0. The exact value of $\bar{A}$ is unknown, but from a statistical error analysis we know that

$$\bar{A} = E[A - cB], \quad (27)$$

where A and B are symmetric matrixes we can compute from image data while c is an unknown constant characterizing the behavior of image noise. Matrixes A and B are random variables since they are computed from data, while c has a definite value determined by the statistical model of noise. Hence, if we put

$$\hat{A} = A - cB, \quad (28)$$

and if we can choose c such that $E[\hat{A}] = \bar{A}$, the unit eigenvector u_m of $\hat{A}$ for the smallest eigenvalue is an unbiased estimate of $\bar{u}_m$. However, we cannot do this unless we know the noise characteristics. On the other hand, if $E[\hat{A}] = \bar{A}$, then

$$E[(\bar{u}_m, \hat{A}\bar{u}_m)] = (\bar{u}_m, E[\hat{A}]\bar{u}_m) = (\bar{u}_m, \bar{A}\bar{u}_m) = 0. \quad (29)$$

So, we attempt to compute a unit vector u_m such that $(u_m, \hat{A}u_m) = 0$ in each iteration step. If $(u_m, \hat{A}u_m) \neq 0$ for the current estimate u_m and c, we define

$$\hat{A}' = \hat{A} - \frac{(u_m, \hat{A}u_m)}{(u_m, Bu_m)}B. \quad (30)$$

Then, $(u_m, \hat{A}'u_m) = 0$. Note that $(u_m, \hat{A}'u_m)$ equals the smallest eigenvalue λ_m of $\hat{A}$.

If u_m and c are the converged values, we have $\lambda_m = u_m, \hat{A}u_m)$ = 0. This does not necessarily ensure that u_m coincides with $\bar{u}_m$. In other words, $(u_m, \bar{A}u_m)$ is not necessarily 0, because the current image data may not necessarily be typical (i.e., may not be a good representative of the statistical ensemble). However, we can expect that u_m is a good approximation with a high probability.

Let us examine the speed of convergence of renormalization. If u_m is the unit eigenvector of the current $\hat{A}$ for the smallest eigenvalue λ_m, matrix $\hat{A}$ is updated at the next step to

$$\hat{A}' = \hat{A} - \frac{\lambda_m}{(u_m, Bu_m)}B. \quad (31)$$

According to the perturbation theorem, the smallest eigenvalue λ'_m of $\hat{A}'$ is

$$\begin{aligned} \lambda'_m &= \lambda_m - (u_m, \left(\frac{\lambda_m}{(u_m, Bu_m)}B\right)u_m) \\ &+ O(\frac{\lambda_m}{(u_m, Bu_m)}B)^2 = O({\lambda_m}^2), \end{aligned} \quad (32)$$

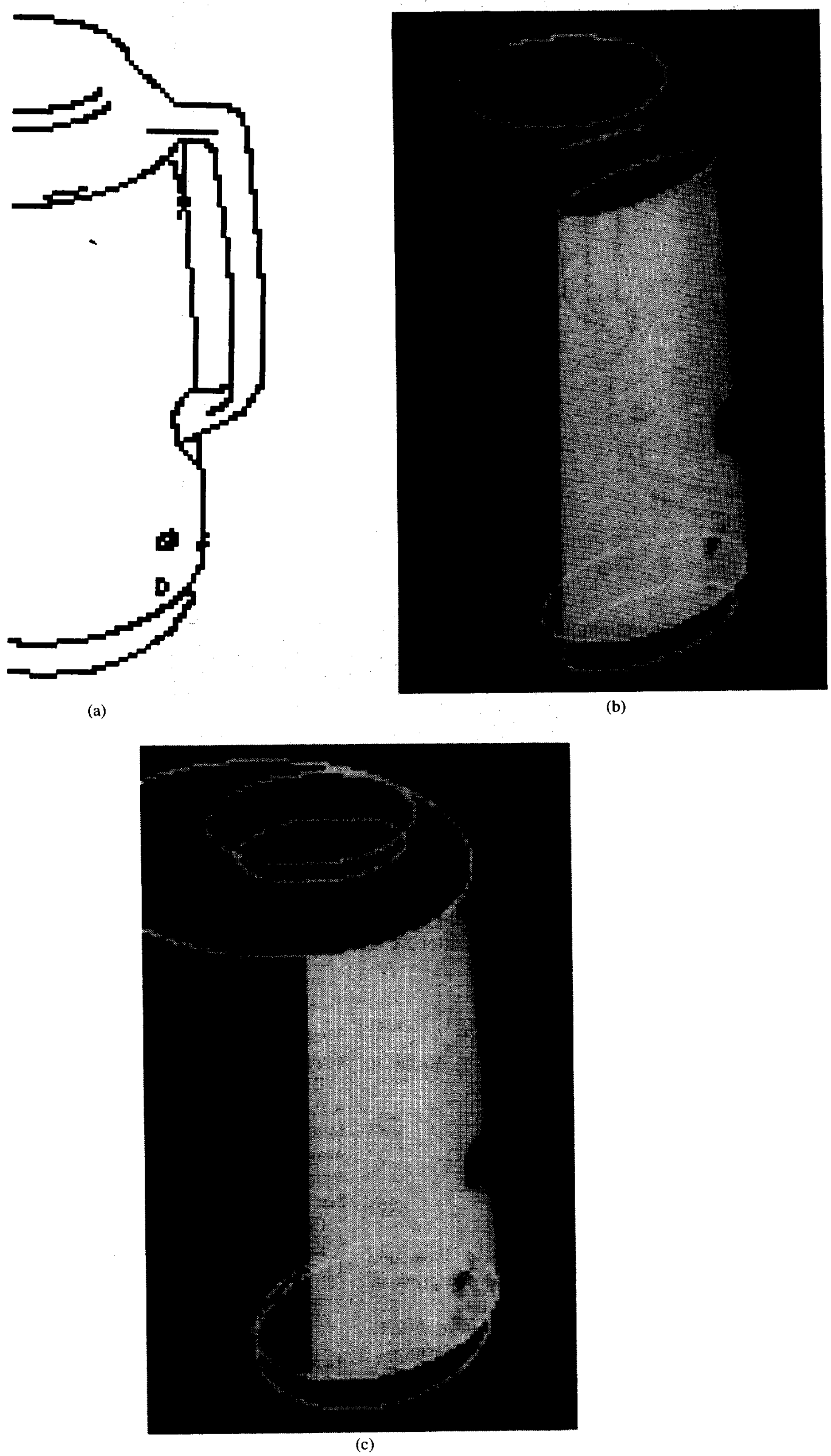


Fig. 4. (a) A half edge image. (b) Least-squares fitting. (c) Renormalization.

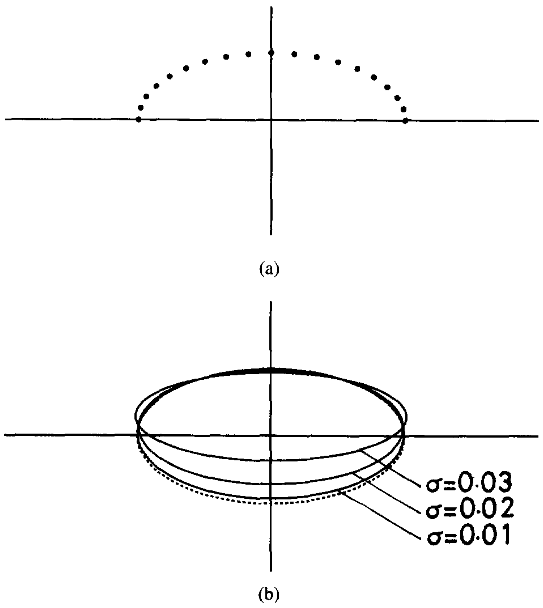


Fig. 5. (a) Points on the upper half of a conic. (b) Theoretically predicted bias.

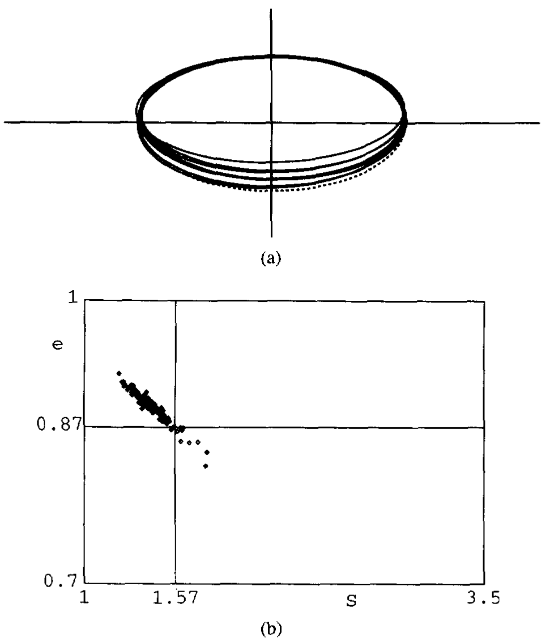


Fig. 6. Conic fitting without renormalization. (a) 10 samples of fitted conics. (b) 100 samples of the area S and the eccentricity e.

where $O(\cdots)^2$ denotes terms of order 2 or higher in $\cdots$. This means that λ_m converges to 0 quadratically.

Let λ_1 and λ_2 be, respectively, the largest and second largest eigenvalues of $\hat{A}$, and u_1 and u_2 the corresponding unit eigenvectors. According to the perturbation theorem, the unit eigenvector u'_m of $\hat{A}'$ at the next step for the smallest eigenvalue λ'_m is

$$u'_m = u_m + \frac{\lambda_m}{(u, Bu)} \sum_{i=1,2} \frac{(u_i, Bu_m)}{\lambda_m - \lambda_i} u_i + O({\lambda_m}^2) = u_m + O(\lambda_m). \tag{33}$$

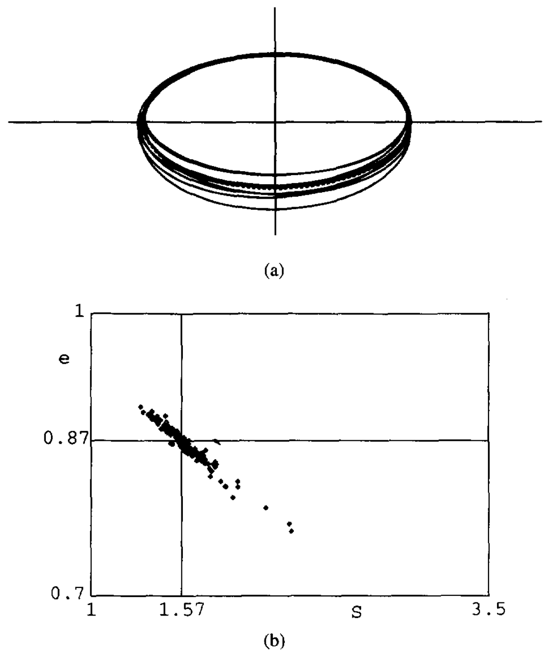


Fig. 7. Conic fitting with renormalization. (a) 10 samples of fitted conics. (b) 100 samples of the area S and the eccentricity e.

If λ_m converges to 0 quadratically, the convergence of u_m is also quadratic. If the optimal weights W_α are computed by using the current eigenvector u_m, the convergence of the W_α is no longer quadratic. However, the convergence is very rapid, and three or four iterations are sufficient for most cases.

Although convergence of renormalization is very rapid, it should be emphasized that the converged values are not necessarily the *exact* values, because the noise is random and unpredictable. The purpose of renormalization is to remove statistical bias (not completely, though). It can be generally proved [10] that if each datum has an independent error of root-mean-square magnitude ν and if the number of data is N, the optimal unbiased estimate has an error of root-mean-square magnitude $O(\nu/\sqrt{N})$, which is the lowest bound that can be achieved.

VII. Conclusion

In this correspondence, we pointed out that the least-squares fit is statistically biased even if the weights are optimally chosen. We then presented a new fitting scheme called *renormalization* for computing an unbiased estimate without knowing noise characteristics. We also discussed relationships to existing methods.

As we have pointed out, conic fitting is essentially a process of *statistical inference* for inferring the locations of the true data points. The Kalman filtering [6], [15] is also an effective method of statistical inference. It is a linearized update rule for optimally modifying the current estimate by a linear operation (orthogonal projection, to be precise) each time a new data point is added. Hence, as many iterations as the number of the data points are necessary. In contrast, renormalization is a nonlinear update rule (computing eigenvectors and eigenvalues) for all data points. Hence, the number of iterations is independent of the number of data points; usually, three or four iterations are sufficient. For this advantage, renormalization is expected to become a standard tool for conic fitting in computer vision applications.

References

[1] A. Albano, "Representation of digitized contours in terms of conic arcs and straight-line segments," *Comput. Graphics Image Processing*, vol. 3, pp. 23–33, 1974.

[2] R. M. Bolle and B. C. Vemuri, "On three-dimensional surface reconstruction methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 1, pp. 1–13, 1991.

[3] F. L. Bookstein, "Fitting conic sections to scattered data," *Comput. Graphics Image Processing*, vol. 9, pp. 56–71, 1979.

[4] D. R. Cooper and N. Yalabik, "On the computational cost of approximating and recognizing noise-perturbed straight lines and quadratic arcs in the plane," *IEEE Trans. Comput.*, vol. 25, pp. 1020–1032, 1976.

[5] E. R. Davies, "Finding ellipses using the generalised Hough transform," *Pattern Recogn. Lett.*, vol. 9, pp. 87–96, 1989.

[6] T. Ellis, A. Abboot, and B. Brillault, "Ellipse detection and matching with uncertainty," *Image Vision Comput.*, vol. 10, no. 5, pp. 271–276, 1992.

[7] D. Forsyth *et al.*, "Invariant descriptors for 3-D object recognition and pose," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 10, pp. 971–991, 1991.

[8] R. Gnanadesikan, *Methods for Statistical Data Analysis of Multivariate Observations*. New York: Wiley, 1977.

[9] K. Kanatani, "Computational projective geometry," *CVGIP: Image Understanding*, vol. 54, no. 3, pp. 333–348, 1991.

[10] K. Kanatani, *Geometric Computation for Machine Vision*. Oxford, UK: Oxford University Press, 1993.

[11] K. Kanatani, "Unbiased estimation and statistical analysis of 3-D rigid motion from two views," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 1, pp. 37–50, 1993.

[12] K. Kanatani and W. Liu, "3-D interpretation of conics and orthogonality," *CVGIP: Image Understanding*, vol. 58, no. 3, pp. 286–301..

[13] Y. Nakagawa and A. Rosenfeld, "A note on polygonal and elliptical approximation of mechanical parts," *Pattern Recogn.*, vol. 11, pp. 133–142, 1979.

[14] K. A. Paton, "Conic sections in chromosome analysis," *Pattern Recogn.*, vol. 2, pp. 39–51, 1970.

[15] J. Porrill, "Fitting ellipses and predicting confidence envelopes using a bias corrected Kalman filter," *Image Vision Comput.*, vol. 8, no. 1, pp. 37–41, 1990.

[16] V. Pratt, "Direct least-squares fitting of algebraic surfaces," *Comput. Graphics*, vol. 21, no. 4, pp. 145–152, 1987.

[17] C. A. Rothwell, A. Zisserman, C. I. Marinos, D. A. Forsyth, and J. L. Mundy, "Relative motion and pose from arbitrary plane curves," *Image Vision Comput.*, vol. 10, no. 4, pp. 250–262, 1992.

[18] R. Safaee-Rad, I. Tchoukanov, B. Benhabib, and K. C. Smith, "Accurate parameter estimation of quadratic curves from grey-level images," *CVGIP: Image Understanding*, vol. 54, no. 2, pp. 259–274, 1991.

[19] R. Safaee-Rad, I. Tchoukanov, K. C. Smith, and B. Benhabib, "Constraints on quadratic-curved features under perspective projection," *Image Vision Comput.*, vol. 19, no. 8, pp. 532–548, 1992.

[20] ——, "Three-dimensional location estimation of circular features for machine vision," *IEEE Trans. Robotics Automat.*, vol. 8, no. 5, pp. 624–640, 1992.

[21] P. D. Sampson, "Fitting conic sections to 'very scattered' data: An iterative refinement of the Bookstein algorithm," *Comput. Graphics Image Processing*, vol. 18, pp. 97–108, 1982.

[22] J. G. Semple and G. T. Kneebone, *Algebraic Projective Geometry*. Oxford, UK: Clarendon Press, 1952 (reprinted 1979).